DOCUMENT RESUME

ED 079 717 CS 000 679

AUTHOR Hall, James W.
TITLE The Development of Memory-Encoding Processes. Final Report.
INSTITUTION Northwestern Univ., Evanston, Ill.
SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
BUREAU NO BR-1-0522
PUB DATE 25 Aug 73
GRANT OEG-5-71-0053(508)
NOTE 58p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Cognitive Processes; *Concept Formation; Elementary Grades; Kindergarten Children; Learning; *Learning Theories; *Memory; Preschool Children; *Recall (Psychological); Verbal Learning

ABSTRACT

Eight experiments are described in this report. In experiment 1, four and five year olds were presented four verbal discrimination (VD) lists. For three of the lists all correct items were either conceptually, acoustically, or affectively similar. The fourth was a control list. The conceptual list facilitated VD learning. Experiment 2 indicated that recognition memory (RM) by kindergartners for target items was better when those items had been preceded by associates than when they had not. In experiment 3, three and four year olds were given extensive paired-associate (PA) training in which response terms were highly familiar words and stimulus terms were words semantically related to response terms but unfamiliar to the subjects. The feasibility of using an entirely verbal PA task with children was demonstrated. Experiment 4 examined the effects of the categorizability of words to be remembered on free recall (FR) of children three through six. In experiment 5 preschool and young elementary school children were given five FR study-test trials with a nine-word list. Experiment 6 consisted of the collection and compilation of word association data for preschoolers. Experiments 7 and 8 examined the modifiability of memory encoding operations of preschoolers and used training exercises to increase the children's awareness of conceptual relationships. (WR)

ED 079717

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Final Report

Project No. 1-0522
Grant No. OEG-5-71-0053(508)

The Development of Memory-Encoding Processes

James W. Hall

Northwestern University

Evanston, Illinois 60201

August 25, 1973

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

CS 000 679

Table of Contents

Acknowledgements 1

Summary 2

Experiments I and II: Young Children's Memory Encoding Reflected in Verbal Discrimination Learning and Recognition Memory Performance 4

Experiment III: Experimentally Acquired Associations in Memory Encoding by Preschoolers 9

Experiment IV: Free Recall of Conceptually Related Words by Young Children 14

Experiment V: Effects of Imagery Instructions on Young Children's Free Recall 18

Experiment VI: Word Associations of Preschoolers 22

Experiments VII and VIII: Modifying Preschoolers' Use of Conceptual Encoding Processes: Two Learning-to-learn Experiments 34

Tables

Table 1: Stimulus and Response Terms Used in Paired Associate Training 11

Table 2: Words Used in the Construction of Free Recall Lists for Experiment IV 17

Table 3: Stimulus Words Used in Experiment VI 23

Table 4: Frequency of Associative Responses by Preschoolers to 44 Common Words 26

Table 5: PAL -- Facilitation Word List Used in Experiment VII 38

Table 6: Conceptually Related Word Lists Used in Free Recall Task in Experiment VII 40

Table 7: Summary of Performance on Test Tasks in Experiment VII 42

Table 8: Mean Performance on Test Tasks in Experiment VIII 50

Figure 1: Mean Number of Words Recalled as a Function of Age, Instructions, and Trials 21

Acknowledgements

The author is especially grateful to Sharon Madsen, who supervised the conduct of Experiments VII and VIII and assisted in the preparation of this report. The assistance and cooperation of Foorman and Susan Mueller, owners and directors of the Tom Thumb Nursery School in Evanston, were invaluable in carrying out Experiments III through VIII. Thanks are due the children and staff of the school, and especially to Gay Hall. The cooperation of the Evanston public elementary schools and the Evanston Child Care Center also was appreciated.

Kent Elwood (Experiments VII and VIII), Marge Goldman (report preparation), Marcia Halperin (Experiments I-III), Gail Page (Experiment VI), and Lannie Wilson (Experiment V) contributed substantially to the research reported here. Others who assisted with the research include Frances Andriacchi, Claudia Brice, Doreen Dolloff, Sandi Eng, Maryann Fritz, Karen Glinert, Susan Hendricson, Donna Kuhn, Kathy Myer, Nancy Klastorin Narron, Marlene Peiser, Mary Ann Romeo, Jean Slater, and Jill Zurlin.

Summary

Eight experiments are described in this report. In Experiment I, four- and five-year-old children were presented four verbal discrimination (VD) lists. For three of the lists all correct items were either conceptually, acoustically, or affectively similar. The fourth was a control list. Only the conceptual list facilitated VD learning, suggesting heavy use of the verbal associative attribute in memory encoding, with no evidence for the involvement of the acoustic or the affective attribute. Experiment II indicated that recognition memory (RM) by kindergarteners for target items was better when those items had been preceded by associates than if they had not. Both the VD and the RM tasks appear promising for developmental research concerning memory encoding.

In Experiment III, three- and four-year-old children were given extensive paired-associate (PA) training in which response terms were highly familiar words and stimulus terms were words semantically related to response terms but relatively unfamiliar to the Ss. Subsequent recognition memory (RM) performance indicated, by an analysis of false recognitions, that the newly acquired associations functioned during the encoding of the RM items, i.e., the PA response terms occurred as implicit associative responses to the PA stimulus terms presented for learning. The feasibility of using an entirely verbal PA task with children as young as three years also was demonstrated.

Experiment IV examined the effects of the categorizability of to-be-remembered words on free recall (FR) of children three through six years of age. Five trials were given, using one of three types of nine-item word lists. The blocked-categorizable list consisted of three instances of each of three categories, and instances were blocked together within the list. The unblocked categorizable list was similar except that instances within categories were distributed rather than blocked. The uncategorizable list consisted of nine conceptually unrelated words. A triple interaction of Age X List Type X Trials reflected a tendency of the older children, but not the younger ones, to recall more from the categorizable lists than from the uncategorizable list, and on later trials only. For the older children FR was best for the blocked categorizable list and words for the uncategorizable list, a result similar to that found in other studies with older children. No such difference occurred for the preschoolers, suggesting that these children have not yet acquired operations or strategies that allow them to profit from the conceptual relations among to-be-remembered items.

In Experiment V preschool and young elementary school children were given five FR study-test trials with a nine-word list. Half the Ss were told to produce visual images during study, and the remainder simply to repeat the words aloud. A facilitative effect of imagery instructions on recall occurred, and that effect appeared to increase with age, consistent with Piaget's analysis of imagery development. Recall also increased markedly with age and over trials.

Experiment VI was not an experiment in the usual sense, but rather consisted of the collection and compilation of word association data for preschoolers. A set of 44 common words was given as stimulus items in a word association test to 48 three-year-old and 48 four-year-old children. There were 24 males and 24 females at each age level. Although the task was an extremely difficult one for the children, data were obtained and are reported here that may be useful to experimenters working with children of those ages.

Experiments VII and VIII examined the modifiability of memory-encoding operations of preschoolers. In Experiment VII preschoolers were randomly assigned to an experimental and a control group, and the experimental Ss were given 21 sessions of training extending over a period of about three months. The training exercises, cast in the form of games, were designed to increase the extent to which the children became aware of conceptual relationships among items and produced category labels to individual instances during encoding and retrieval. At the conclusion of training, a series of "tests" were administered to examine the effect of training. No clear cut effects of training were found and Experiment VIII was then conducted. Experiment VIII was similar to Experiment VII, except that training and testing both were more extensive. The results indicated considerable learning of a relatively specific nature occurred, but relatively little by way of more generalized training effects. In particular, in neither experiment did the experimental Ss exceed the control Ss in their FR performance with categorizable lists despite evidence that the experimental Ss were more likely to produce the appropriate category labels in such a task. Implications of these results were considered.

Experiments I and II

Young Children's Memory Encoding Reflected in Verbal Discrimination Learning and Recognition Memory Performance*

The experiments reported here constitute an attempt to develop additional techniques for use in examining memory-encoding processes, and particularly in tracing the development of such processes. A number of previous studies have employed a false recognition technique to attack such questions. In this approach the frequency of false recognitions of words bearing a particular relationship to a given set of to-be-remembered items is compared with a baseline frequency of false recognitions of items unrelated to previous words. The most common relationship examined has been a verbal associative one. Thus, if an associate (e.g., CHAIR) of a to-be-remembered item presented earlier (e.g., TABLE) is incorrectly judged "old" more frequently than some equally familiar unrelated word (e.g., GRASS), it is inferred that the verbal associative attribute was active in the encoding of the item. In the above example it is inferred that CHAIR had occurred as an implicit associative response (IAR) in the encoding of TABLE. This technique has been useful in demonstrating that both the associative and the acoustic qualities of words are heavily involved in memory encoding both by adults (e.g., Underwood, 1965) and by children (e.g., Bach and Underwood, 1970).

Despite the demonstrated utility of the false recognition technique, it possesses an important limitation where developmental and comparative studies are concerned. Briefly, that limitation is that differences in the degree to which verbal associative responses occur during encoding may be obscured or distorted by corresponding differences in subjects' abilities to distinguish, during the test phase, between words actually presented and those that had occurred only implicitly during the encoding of the presented item. This problem was illustrated in two experiments (Hall and Ware, 1968, Exp. II; Hall, 1968) in which IAR-produced false recognitions were more frequent by kindergarten than by third-grade children. Rather than ascribing these findings to a decrease in IAR occurrence with age, it seemed more likely that the older children simply were better able to discriminate as described above. Evidence that such discrimination does occur has been reported previously (Hall, Sekuler, and Cushman, 1969) and the developmental implications have been discussed at some length (Hall, 1969).

*The results of these two experiments are described in an article of the same title by James W. Hall in Psychonomic Science, 1971, 25 (2), 91-93.

In the experiments reported here two standard laboratory tasks were adapted for use in examining the role of various word-attributes in memory encoding by young children as alternatives to the false-recognition technique just described. In Experiment I a verbal discrimination task is employed to study the role of the verbal associative, the acoustic, and the affective attribute in memory encoding by 4- and 5-year olds. Experiment II employs the recognition-memory technique, but focuses on correct rather than false recognitions.

Experiment I

In Experiment I, four verbal discrimination lists were presented to 4- and 5-year olds. The lists differed both in the specific words used and in the relationships among correct items from pair to pair within each list. In one list all correct items rhymed with one another; in a second each was a member of the same common category; in the third each was a word judged to possess positive affective qualities; and in the fourth (the control list) no systematic relations existed among them. It was assumed that to the extent that acoustic, conceptual, and affective qualities of the words were involved in encoding there would be a corresponding facilitation of learning for the respective lists relative to the control list. Following Underwood's (1969) view of the acoustic attribute as a more primitive one, it was anticipated that with the rhyming list considerable facilitation would occur with even the youngest children. The verbal associative attribute, considered more "advanced," was expected to increase in prominence with age in relation to the acoustic attribute. No specific prediction was made in the case of the affectively similar items.

Method

Subjects. The Ss employed were 16 children of ages 4:4 to 5:3 years (M = 4:10) and 16 children of ages 5:4 to 6:3 years (M = 5:9) enrolled in a day-care center in Evanston.

Lists and procedure. Four verbal discrimination lists were employed. The Conceptual list consisted of seven pairs in which the correct items were common zoo animals (e.g., LION, MONKEY) frequently encountered in childrens' stories and highly familiar (at least at a verbal level). In the Acoustic list the correct items consisted of seven common one-syllable words rhyming with SUN. In the Affective list the correct items were comprised of seven common words each judged to have positive affective value for children and to be relatively unrelated associatively to one another (e.g., LOVE, FRIEND). The incorrect items in the above list, and all items on the Control list were familiar words that appeared to be unrelated to one another in any systematic way. In all lists the words within each pair were identical in number of syllables and in form class.

Since each S was exposed to every list, a counterbalancing procedure was used resulting in four orders of presentation. Approximately equal numbers of Ss at each age level received each presentation order. Each S received each list in a separate experimental session with either two or three days intervening between sessions. At the initial session the S was told that he would play some word games, that the games would take several days, and that he would receive a prize if he "tried hard" in the games. He was told that words were to be spoken "two at a time," that one word would be the right one, and that he was to try to remember the right words. The experimenter then pronounced each pair at the rate of approximately 4 seconds per pair, with the correct item indicated as follows: "SHOE and QUEEN, SHOE is the right word." After presenting the pairs this way, two test trials were given in which the S was presented with each pair in the following manner: "Which is the right word, QUEEN or SHOE?" Since feedback was given on these trials, they functioned also as learning trials. The order of presentation of pairs and of words within pairs varied randomly over the three presentations, but was constant for all subjects.

Results

An Age X List X Trials analysis of variance indicated a reliable main effect of age [F (1,30) = 6.02, $p < .05$] and borderline reliability for the main effects of Lists [F (3,30) = 2.19, $p < .10$] and Trials [F (1,30) = 3.56, $p < .10$]. No interactions approached significance. Collapsing over lists and trials, the mean number of correct responses was 4.75 (SD = 1.23) for the younger group and 5.74 (SD = 1.05) for the older group. Thus the older Ss exceeded the younger ones by an average of one word per S. Collapsing over lists and ages, the mean number of correct responses was 5.15 (SD = 1.17) for Trial 1 and 5.34 (SD = 1.36) for Trial 2, a difference that, while of borderline significance, shows relatively little learning. In fact, although the relevant interaction was nonsignificant, it is interesting that there was substantial improvement from Trial 1 to Trial 2 for the Conceptual and Control lists, and none at all for the Acoustic and Affective lists.

Of major interest in the experiment was the List's effect. The relevant means, collapsing over ages and trials, were 5.62 (SD = 1.54) for the Conceptual, 5.17 (SD = 1.55) for the Acoustic, 5.14 (SD = 1.60) for the Affective, and 5.05 (SD = 1.27) for the Control. Only the Conceptual list was learned reliably better than the Control list [t (31) = 2.76, $p < .05$, following Dunnett's procedure as described by Winer, 1962]. There was not even the suggestion of facilitation resulting from either the acoustic similarity or the positive affective quality.

Experiment II

The disadvantage of the false-recognition procedure, wherein the S is working (often successfully) to avoid errors has been discussed above. Experiment II, therefore, was designed to explore the effectiveness of an alternative recognition-memory procedure wherein the occurrence of IARs would be inferred from an increase in correct recognitions rather than in false recognitions. In this procedure, target items were preceded in a list of to-be-remembered items by words with which the target items are associated. It was assumed, then, that if the target item occurs as an IAR in addition to its explicit occurrence in the list, the probability of correct recognition of the target item when it is presented subsequently will be increased relatively to that probability for a word that has occurred only explicitly.

Method

Subjects. The Ss were 18 kindergarten children (12 boys and 6 girls) of ages 5:4 to 6:4 years (M = 5:10) enrolled in a public elementary school in Evanston.

Lists and procedure. A list of 24 to-be-remembered items (List 1) was presented aurally to each S, followed, after a 10-second interval, by the presentation of the 15-item test list (List 2). Two forms of List 1 were prepared, with half the Ss randomly assigned to each form. Each form consisted of four Target items, immediately preceded by three Target Associates and immediately followed by a Control item (an item unassociated with other items on the list). The Target items on one form, along with their Target Associates were HORSE (SADDLE, GALLOP, PONY), WATER (DRINK, THIRSTY, RAIN), SLEEP (BED, DREAM, NAP), and FRUIT (BANANA, ORANGE, APPLE). On the other form the Target items and Target Associates were DOG (PUPPY, BARK, PET), CUT (KNIFE, SLICE, SCISSORS), FAST (QUICK, SLOW, RUN), and BIRD (ROBIN, FEATHERS, EAGLE). The Control items were DOG, CUT, FAST, and BIRD for the first form and HORSE, WATER, SLEEP, and FRUIT for the other. Thus, each Target item of one form served as a Control item on the other form, a counterbalancing procedure that eliminates the possibility of biased results due to particular properties of items chosen as target or as control items. The remaining four items in List 1 were fillers, two occurring at the beginning and two at the end of the list. Each word in List 1 was pronounced twice by the experimenter with a 3-second interval between words. The Ss were asked to listen carefully and try to remember the words.

List 2 was the same for all Ss and consisted of the eight Target-Control items named above, and four "new" items unrelated to any previous ones. The order of these 12 items was determined

randomly and they were preceded by 3 filler items which were used to insure that the Ss understood the task. The Ss were told, "Now I'm going to see whether you can remember the words that I said to you earlier. You listen carefully and don't let me trick you. Ready?" The first test-list item (a filler item) was then spoken as follows: "Did I say TELEPHONE?" If the S correctly said "No," he was praised, and if he said "Yes" he was corrected. The remaining items were presented in the same fashion, but feedback was given only for the filler items. If a S did not respond immediately, or if he showed uncertainty, the question was repeated and he was told to guess whether or not the item had been presented previously. All responses were recorded by the E, who then went immediately to the next item.

Results

The mean number of correct recognitions was 3.06 for Target items and 2.28 for Controls. Expressed in percentages, 76% of the Target words and 57% of the Control words were correctly identified, whereas only 4% of the New words were incorrectly judged old. The mean Target-Control difference of .78 was highly reliable [t (17) = 3.30, $p < .01$], indicating facilitation of learning of the Target items produced by the presence of the Target Associates.

Discussion

Both the tasks employed in the experiments just reported appear to be promising insofar as techniques for the study of developments in memory-encoding processes in young children is concerned. No Ss were lost in either experiment, and both yielded data of interest concerning the memory-encoding characteristics of the children.

The results of Experiment I were of particular interest in that whereas the verbal associative characteristics of items were clearly operative during learning, no evidence that the acoustic attribute was involved was found. This appears surprising in light of the notion of acoustic encoding as a primitive process which becomes decreasingly prominent relative to the verbal associative attribute as the child develops (Underwood, 1969). It is interesting that when several experienced nursery school teachers were asked to predict the outcome of Experiment I, all of them anticipated marked facilitation by the rhyming relationship, most predicting as great or greater a facilitative effect as for the Conceptual list. Surely this finding deserves further investigation, since on the surface it suggests that some revision of thinking about the early role of the acoustic attribute in memory-encoding may be necessary. A hypothesis that seems worth investigating is that the retention of acoustic information about a word requires, or is facilitated by, the pronunciation (overtly or

covertly) of the word. If this is true, the lack of such retention in very young children is more understandable, since there is ample evidence that they are less likely than are older children to rehearse spontaneously items presented for learning (e.g., Flavell, Beach, and Chinsky, 1966). In any case, the data suggest that the verbal discrimination task may be a useful one in attacking developmental questions concerning memory-encoding processes.

The recognition-memory task, as it was used in Experiment II, also appears promising. The data suggest that verbal associations are heavily involved in the memory-encoding of children as young as 5 years, and thus are consistent with the results of previous studies employing the false recognition technique (Hall & Ware, 1968; Hall, 1968), and with the findings of Experiment I with respect to the facilitation observed for the Conceptual list.

Experiment III

Experimentally Acquired Associations in Memory Encoding by Preschoolers*

The present experiment examines the extent to which very young children spontaneously produce newly acquired associations during memory encoding. The general question of concern in this paper is whether the observed increase with age in the use of verbal associations during memory encoding is entirely the result of increased availability of associations or is due also to the development of some "spontaneous production mechanism." Evidence that the occurrence of verbal associations during encoding does increase over an age range of about seven to twelve years has been furnish by Bach and Underwood (1970) and by Felzen and Anisfeld (1970). In these studies the frequency of false recognitions of natural language associates of previously presented words increased with age, presumably due to an increase with experience in the strength of these associations However, it appears possible that such differences might be due in part at least to an IAR production deficiency described by Flavell, Beach, and Chinsky (1966) wherein very young children less frequently produced verbal mediators spontaneously during learning although such mediators were produced on demand. Three recognition memory (RM) experiments recently conducted with three- to five-year old children provided no evidence of such deficiency in the case of IAR production during learning. For example, in one of these experiments (Hall & Halperin, 1972, Exp. II) three-, four-, and five-year-olds were presented a series of highly familiar words to be remembered, followed by a test list that included words associatively and conceptually related to certain of the to-be-remembered items. False recognitions

*This experiment was reported in an article of the same title by James W. Hall in Psychonomic Science, 1972, 29 (2), 113.

were more frequent for the associatively and conceptually related items than for unrelated control items at every age level and no increase with age in this respect was noted. It appeared, then, that even children as young as three years of age spontaneously produce IARs much as do older children and adults, provided that the relevant associations have been acquired, i.e., provided that highly familiar experimental material is employed.

A primary purpose of the present experiment was to provide further evidence concerning the question raised above by examining the extent to which three-year-old children spontaneously produce associates newly acquired in the laboratory when encoding words for memory. A second experimental objective was a methodological one. Extensive use of standard verbal learning tasks with very young children is a relatively recent development. It is clear that such efforts are beset with difficulties having to do mainly with obtaining and maintaining the attention of Ss and posing instructions in such a way that Ss understand the tasks. We have found it necessary often to conduct special pre-training sessions, to use brief multiple sessions, and to explore various procedures through considerable pilot work. The major methodological objective of the present experiment concerned the feasibility of employing a purely verbal paired associate (PA) task involving multiple sessions with children as young as three years. The idea here was first to "teach" a set of new associations to the children through PA training, then to determine, through the use of a RM task, to what extent these new associations would occur spontaneously during memory encoding.

Method

Subjects

The Ss consisted of 14 boys from two years, 10 months to four years, one month (mean = three years, seven months) and 14 girls from two years, 11 months to four years, four months (mean = three years, six months) who were enrolled in a private nursery school in Evanston. The limited SES data available indicated that the families of these children were primarily lower-middle class with some working-class and upper level professionals as well.

Materials and Procedure

Phase I: PA training. The items used in PA training are shown in Table 1. Every S received the same 12 response words, each of which was judged to be highly familiar to the Ss prior to commencing PA training. For seven boys and seven girls the words in List 1a served as stimulus terms, whereas the words in List 1b served as stimulus terms for the remaining Ss. As may be seen in Table 1,

Table 1

Stimulus and Response Terms Used
in Paired Associate Training

Stimuli		
List 1a	List 1b	Responses
BEAGLE	POODLE	DOG
RAPID	SPEEDY	FAST
POND	BEACH	LAKE
COTTAGE	CABIN	HOUSE
TACK	SPIKE	NAIL
HOG	PORK	PIG
BROOK	STREAM	RIVER
YACHT	TUG	BOAT
SLACKS	TROUSERS	PANTS
CHEERFUL	JOYFUL	HAPPY
CROW	SPARROW	BIRD
WOOL	LAMB	SHEEP

the stimulus terms were either subordinates of their corresponding response terms or were highly related semantically (e.g., synonyms). Since it was important that the association between pair members be relatively low at the outset of training, an oral pretest was given in which Ss were asked what each stimulus word meant. The results confirmed the desired low association, since most Ss indicated lack of familiarity with all of the stimulus terms.

Six of the 12 pairs were introduced in Session 1 with a study trial of the following type: "TUG. A TUG is a kind of BOAT. When I say TUG, you say BOAT." Four anticipation trials then were given, and the session was terminated. In Session 2 the remaining six pairs were introduced. The procedure in that session was the same as in Session 1 except that following the four anticipation trials on the second set of six pairs, two anticipation trials were given with the items introduced in Session 1.

The remainder of the PA training consisted of seven sessions of four anticipation trials per session, except that (a) if a S performed perfectly on two consecutive trials within a session that session was terminated, and (b) if a S performed perfectly in two consecutive sessions the sessions were terminated. In such cases the Ss were assigned a perfect score for those trials and/or sessions that were omitted.

The order of pairs presentation was varied randomly both within and between sessions, and the same order was used for all Ss receiving a particular set of pairs. Presentation rate varied somewhat according to the idiosyncracies of the Ss, but in general was such that sessions typically ran from five to eight minutes in duration.

Phase II: The RM task. The RM task was administered first 7-10 days following completion of the PA training. It was administered in two parts in order to reduce attentional problems. For each part, 10 items were read in a predetermined random order at a 4-5 second rate with Ss instructed to pay careful attention and try to remember the words. After a break of about 30 seconds the test list was administered using the following format: "BOAT. Was BOAT one of the words I said to you?" The two parts were administered during a single session and were separated by a break of about 30 seconds.

Among the total of 20 to-be-remembered items, randomly ordered for presentation, were 12 that had served as stimulus words for PA training. Of these 12 items, six were from PA list 1a and six from PA list 1b, i.e., only six had served as PA stimulus items for any given S. The remaining eight to-be-remembered items also appeared in the test list, and are termed Repeated (R) words. In addition to these eight R words, the test list included the 12 PA response words, and 12 words

that were unrelated to any of the to-be-remembered items (termed Unrelated Control or UC words). Note that for any given S only six of the 12 PA response terms included in the RM test had been paired with any of the to-be-remembered items in the RM task. These six words are termed Experimental Associates (EA words) whereas the remaining (equally familiar) PA response terms are termed Control Associates (CA words), in that it is the difference in frequency of false recognitions to these two classes of words that would reveal the effects of PA training on IAR production during memory encoding.

Finally, since all Ss continued to be available for experimentation for a period of several months, the recognition memory task was again administered some seven to eight weeks following PA training to examine the persistence of effects found in the earlier testing.

Results and Discussion

PA Training

Two Ss were eliminated from the study due to their consistent failure to follow the PA task instructions, and a third because of frequent nonattendance during the period of the study. For the remaining 25 Ss performance increased monotonically across sessions for Sessions 3 through 9 from a mean of 24.44 (SD = 12.55) items correct over the four trials in a session (of a possible 48) to a mean of 41.72 (SD = 5.67) or about 87 per cent correct during the final session. These data suggest that the task was well suited to the capabilities of the Ss, none of whom had had previous experience with such a task.

It is clear from the above results that the use of a purely verbal paired associate task with children as young as three years of age is entirely feasible. No previous demonstration of this feasibility was found in a search of available literature. It was necessary, of course, to modify considerably the standard laboratory procedures by introducing the full set of twelve pairs gradually and by employing a series of brief sessions rather than massing the trials in one or two longer sessions. In addition, auditory presentation was required, and the instructions given on the initial study trial were modified to stress the semantic connection between the members of pairs. Nevertheless, the task was a paired associate one and performance did increase steadily over sessions to a relatively high performance level in the final session. Moreover, this was accomplished with Ss who had

no prior associate training and who represented a rather wide range in SES, and with pairs in which stimulus items were relatively low in meaningfulness for these young children.

RM Performance

Of the 25 Ss who completed the PA training, 13 (six boys and seven girls) had received PA List 1a and 12 (six boys and six girls) had received PA List 1b. To insure the effectiveness of the PA counterbalancing procedure, the data for one of the seven girls who received PA List 1a (randomly determined) were discarded so that the analyses of the RM data are based on 24 Ss, six in each of the counterbalancing conditions.

In the first RM session the mean percentages of judgments of "old" were 84 for R words, 58 for EA words, 38 for CA words, and 34 for UC words. Major interest was in a comparison of EA and CA words in frequency of false recognitions (i.e., judgments of "old"). That mean difference of 1.16 (SD = 1.81) was highly significant, t (23) = 3.09, $p < .01$. Note that the CA and UC items differed only slightly and nonsignificantly in that respect.

In the second RM session the mean percentages of "old" judgments were 82 for R words, 33 for EA words, 21 for CA words, and 13 for UC words. The mean difference between EA and CA words of .67 (SD = 1.72) was of borderline reliability, t (23) = 1.85, $.10 > p > .05$. Again, the difference between the CA and UC items was not significant.

These data indicate that after one week and probably also 7-8 weeks following PA training the newly acquired associations functioned during the encoding of the RM items. Such results offer no support for the notion of an IAR production deficiency as introduced at the outset of this paper. However, the present data do not preclude such a possibility. That is, it is possible that the acquisition of new associations by older children would have resulted in even greater use of them for memory encoding. In any case, it appears that newly acquired verbal associations are readily activated during memory encoding by children as young as three years.

Experiment IV

Free Recall of Conceptually Related Words by Young Children

On a free recall (FR) task the performance of normal adults is markedly facilitated when items can be categorized easily (e.g., MONKEY, LION, GIRAFFE) or when natural language associations exist

among the list items (e.g., SCISSORS, CUT). A number of investigators have been concerned with the development of such effects in children. Laurence (1967) found that preschoolers (mean CA = 5.0) performed better in free recall of names of pictured objects when all objects were from a single common category (animals or items of clothing) than when they were not obviously related. (Unfortunately interpretation of these findings is complicated by the fact that such a difference was not found for children a year older.) Vaughan (1968) examined free recall of picture labels by first-, fourth-, and seventh-graders using a categorizable-blocked list and an uncategorizable list. Free recall was better for the categorizable-blocked list at every level. (A categorizable-unblocked list was not included.) Nelson (1969), comparing the free recall of words by five- and eight-year olds, found better recall for a categorizable-unblocked list than for an uncategorizable list for the eight-year olds, but not for the younger subjects. Cole, Frankel, and Sharp reported two experiments that are pertinent here (Cole, et al., 1971, Experiments II and III). Both involved the free recall by children of a set of 20 familiar nouns. In Experiment II these nouns consisted of five instances of each of four categories presented either in a blocked or an unblocked fashion to children in grades 1, 3, and 8. Performance improved markedly with age and was better for blocked presentation, but the anticipated interaction between these was not found. In Experiment III the performance of children in grades 1, 4, 6, and 9 was examined with either a categorizable-unblocked or an uncategorizable list. Only a small (though significant) difference between the list conditions was found, and again contrary to expectations, that difference was constant across age. Relatively little clustering was found at any age level, and clustering was nearly absent at grade 1.

In short, the above studies have not clearly demonstrated the expected developmental trend toward increasing facilitation for categorizable versus uncategorizable items. The present study was designed to provide further evidence concerning such a trend by examining effects of item categorizability on multiple-trial FR for children from three through six years of age.

Method

Subjects

The Ss included 27 three-year-olds and 27 four-year-olds enrolled in a private nursery school in Evanston. Also included were 27 kindergarteners and 27 first-graders enrolled in a public elementary school in Evanston.

Design

At each age level Ss were assigned randomly to one of three list conditions: (1) a list of nine unrelated words, (2) nine words consisting of three instances of each of three categories, with items within a particular category blocked during presentation, and (3) the same list as (2) but with items in each category distributed rather than blocked during presentation. Thus, nine Ss served in each list condition at each age level. To prevent confounding of particular words with experimental condition, a pool of 27 words consisting of three words in each of nine categories was generated (see Table 2). A counterbalancing procedure then was used that resulted in an equal number of appearances of each of the 27 words in each of the three list conditions.

Results and Discussion

An Age x List Type x Trials analysis of variance showed highly reliable main effects of both Age and Trials. For that analysis the three- and four-year-olds were combined (the preschool group). The mean recall scores, collapsing over trials and conditions, were 5.81 for the school group and 4.34 for the preschoolers. Collapsing over conditions and age, the mean recall scores increased from 3.57 on trial one to 6.24 on trial five. Two of the interactions also were significant. First, the Age x Trials interaction ($p < .05$) evidently reflected a somewhat sharper increase over trials for the older Ss. The second order interaction (Age x List Type x Trials) also was significant ($p < .05$), and it was that effect that was most interesting in the present data. That interaction reflected the fact that List Type led to differential performance only for the older children, and, further, that this difference appeared only in later trials. Thus, the difference among the means for the three list types was very small (less than .40) both for early and late trials of preschoolers and for early trials of the school children. However, on the final trials of the school children that difference was 1.36, the means being 6.17 for the unrelated list, 6.72 for the related but not blocked list, and 7.53 for the blocked list. In short, amount and arrangement of conceptual relations among to-be-remembered items had no effect with preschoolers, but did influence the kindergarten and first-grade children after some experience with the list (i.e., on later trials).

Generally, it has been predicted that with increasing age children become increasingly able to capitalize on conceptual relationships among items to improve their recall of such items. That prediction was clearly confirmed in the present data for children from three through six years of age. The fact that others have not found clear-cut evidence for a continuation of this trend

Table 2

Words Used in the Construction of
Free Recall Lists for Experiment IV

shirt	window	lake
pants	door	river
hat	wall	ocean
truck	dollar	mouth
bus	penny	ears
car	nickel	nose
knife	yellow	pig
fork	blue	cow
spoon	green	horse

beyond this age range (e.g., Cole, Frankel, & Sharp, 1971) remains somewhat puzzling. It seems highly unlikely that the abilities involved here reach their peak by age six. Nor is it clear from research to date just how the facilitative effect occurs when it does occur. Does this effect reflect activities during encoding, or is it due entirely to recall activity. Finally, to what extent and by what means is it possible for three- and four-year-olds to learn, through training, to capitalize on conceptual relationships for improved performance on memory tasks? The latter question is the subject of Experiments VII and VIII in this report.

Experiment V

Effects of Imagery Instructions on Young Children's Free Recall

There is evidence that instructing adults to produce images of objects denoted by words can facilitate the free recall of those words (Gupton & Frincke, 1970). Whether similar effects obtain with preschool and primary school age children is the general concern of this paper.

Piaget and Inhelder (1969, p. 71) have described the mental images of children in the preoperational period (generally children below six or seven years of age) as "reproductive" and "static." They are reproductive in that the child can produce only images of that which he has perceived previously, and they are static in that movements and transformations are not imageable. In contrast to the reproductive imagery of preoperational children, anticipatory images begin to be mastered by children of about primary school age. According to Piaget and Inhelder (1969, p. 71), anticipatory imagery refers to the production of images of previously unperceived combinations, movement, or transformations of objects.

There is reason to believe that the facilitative effects of imagery on recall are to some degree a function of the extent to which relatively complex anticipatory images are generated. Paivio (1971, p. 329) has cited evidence that suggests that imagery instructions are particularly facilitative of free recall when they result in the production of images of combinations of to-be-remembered items. For example, results of a study by Bower, Lesgold, and Tieman (1969) indicated that adult Ss formed images which included as components the referents of several to-be-remembered words. Moreover, when a regrouping of these items from trial to trial was required, recall suffered.

Assuming, then, that the formation of images in which several items appear is superior to the occurance of single-item images, and assuming that such complex imagery is more difficult (perhaps impossible) for pre-operational children, one would predict at best only a slight facilitative effect of imagery instructions on free recall by preschool as compared to school-age children. Accordingly, the present experiment examines the effects of imagery instructions on multiple trial free recall by four-year old and by six- and seven-year old children.

Method

Subjects

The Ss were 40 nursery school children and 40 first-graders (mean CAs of 4-7 and 6-11 respectively).

Design and Procedure

At each age level the Ss were randomly assigned to an imagery and a control condition with the provision that there be 20 Ss in each condition. The same nine-word list was presented to all Ss in the following randomly determined order on each of five trials: SHIRT, CAR, COW, RIVER, WINDOW, NOSE, DOLLAR, KNIFE, and YELLOW. Prior to the first study trial all Ss were told that their job was to try to remember as many of the words as they could. The Control Ss were told that they should repeat aloud each word as it was presented (orally) to them. The Imagery Ss were told that they should make a picture in their heads of what each word was about. For the Imagery Ss a brief pretraining procedure was then used to ensure that the Ss understood and could comply with the imagery instructions. This procedure involved, for example, asking each child to make a picture in his head of his mother in the kitchen, or his favorite toy, and so on until the E was satisfied that the child was willing and able to comply with the instructions.

The five study-test trials then were given in which the list was read at a five-second rate followed by an untimed test period in which the E recorded all responses. The Ss in the Imagery condition were reminded to form pictures in their heads at the beginning of each study trial.

Results and Discussion

Figure 1 shows the performance of each condition at each age level for each of the five trials. An Age x Instructions (Imagery vs. Control) x Trials analysis of variance showed all three main effects to be significant, F (1, 76) = 55.30 for Age, F (4, 304) = 14.68 for Trials, $p < .00^{1}$ for both, and F (1, 76) = 6.15, $p < .05$

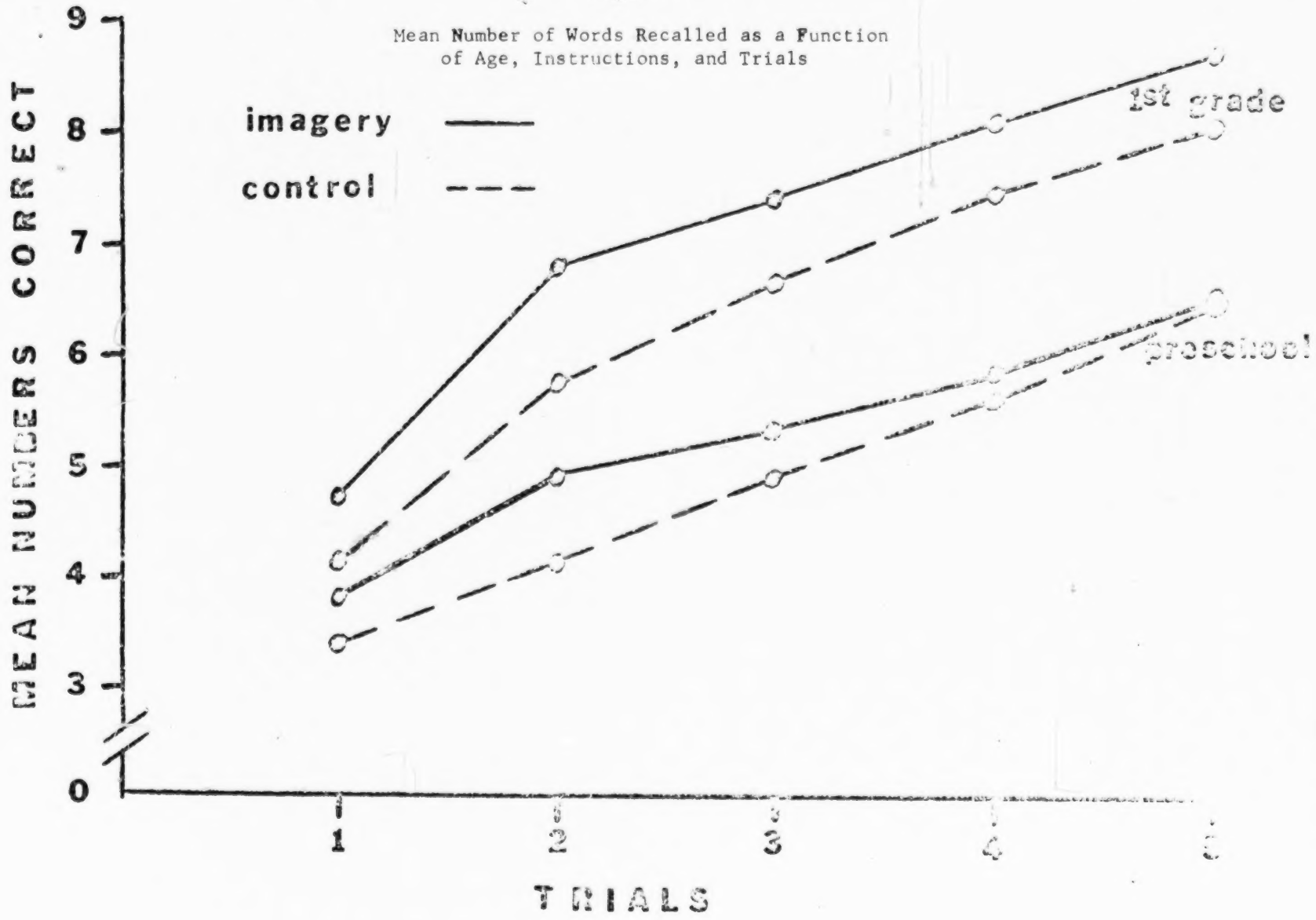
for Instructions. Thus, older Ss were superior to younger ones, performance increased over trials, and, most important, imagery instructions facilitated performance. The only reliable interaction was between Age and Trials, F (4, 304) = 4.87, $p < .01$, evidently mainly reflecting the sharper increase from Trial 1 to Trial 2 for Ss under imagery instructions.

Of particular interest was the possibility of an increase with age in the facilitative effects of imagery instructions. Although the relevant interaction (Age x Instructions) was not reliable, it may be seen from Figure 1 that the difference found was clearly in the predicted direction. In fact, when t tests were applied to the total scores at each age level the results were in line with the expectations. The mean numbers correct for preschoolers were 26.34 (SD = 4.40) for the Imagery condition and 24.50 (SD = 5.30) for the Control condition. The mean difference of 1.95 was not significant, t (38) = 1.26, $p > .10$. The mean recall scores for first-graders were 35.70 (SD = 4.63) and 32.05 (SD = 5.30) for the Imagery and Control conditions respectively. That mean difference of 3.65 was significant, t (38) = 2.26, $p < .05$. The probability of a "real" age difference in effects of imagery instruction is further increased by the likelihood of a ceiling effect with the older children on later trials. For example, as early as Trial Three, 10 of the Imagery Ss obtained scores of eight or nine, whereas only four of the Control Ss obtained such scores on that trial. Thus, it appears likely that the magnitude of the age differences in effects of instruction is underestimated by the data shown in Figure 1.

In short, the evidence points to a facilitative effect of imagery instructions for six- and seven-year olds but is equivocal regarding such an effect for four-year olds. This suggested age difference is in line with expectations based on the assumptions that imagery in children below six or seven years old is deficient in certain respects, and that the qualities absent in the images of such children are among those that facilitate recall. Whether these deficiencies are of the precise nature of those proposed by Piaget (1969, p. 71) remains to be determined. It was the impression of the E that the younger children had more difficulty carrying out the imagery instructions. Thus, despite the pretraining in which all Ss reported forming the appropriate images, it is quite possible that during presentation of the words some of the younger Ss simply did not comply with the instructions as consistently as did the older children.

Figure 1

Mean Number of Words Recalled as a Function of Age, Instructions, and Trials



Experiment VI

Word Associations of Preschoolers

Normative word association data have been available for school age children for several years (Palermo & Jenkins, 1966; Entwisle, 1966), and these data have proven most useful in investigations of children's verbal learning and behavior. Unfortunately, such data have not been published for younger children. Moreover, the developmental changes in associative responses that have been observed during the early school years suggest that the use of these data to estimate response characteristics of younger children would be, at best, risky.

The present study constituted a step in the direction of providing useful normative word association data for preschool children. A second objective of this work was to develop procedures that were effective for the collection of such data. As those who have done such work can attest, obtaining word association data even from kindergarten and first-grade children is not easy. A relatively large number of children that age simply don't seem to "catch on" to the task. It was anticipated that this problem would be even more severe for preschoolers, so considerable effort was devoted to preliminary work to develop procedures that would provide the best possible data.

Method

Subjects

The subjects were 108 children enrolled in nursery school and day care programs in Evanston. From all indications, the socio-economic backgrounds of these children varied widely, with most falling in the lower-middle and upper-lower classes. The responses of 12 subjects were excluded from final compilation on the grounds that these subjects showed no evidence of understanding and/or of willingness to cooperate. That is, these subjects failed to give a single response scored as acceptable according to criteria to be described in the Results section of this report. The final sample, then, consisted of 24 boys and 24 girls at each of ages three and four.

Materials and Procedure

The stimulus items consisted of the 44 words shown in Table 3, administered orally and in the order shown, at approximately a five-second rate. Each subject participated in two sessions with half the words presented at each session. Preliminary work indicated that nearly all these particular words were familiar to

Table 3

Stimulus Words Used in Experiment VI

hammer	fruit
pretty	tell
sad	lion
salt	run
doctor	carry
sleep	mountain
hand	heavy
door	cry
thirsty	listen
needle	ocean
river	wild
boy	lamp
car	flower
sit	sheep
slow	up
cold	quiet
black	bread
bed	loud
table	happy
come	chair
moon	music
gallop	bird

all the children, and that the presentation of 22 words during one session was close to the maximum for many children's ability (or willingness) to sustain attention. Further, while most of the children found the task sufficiently palatable to return for a second session with some urging, a third session seemed out of the question. Nearly all the words used were among those used by Palermo and Jenkins (1966) or by Entwisle (1966) with school-aged children.

Three female experimenters administered the word association test. The instructions given to subjects were as follows:

> I have a word game to play today. Sometimes when you hear a word it makes you think of some other word. For example, if I said the word DOG, you might think of the word CAT, or ANIMAL, or BARK. Now the way we play this game is that I'll say a word and you tell me the first different word that it makes you think of. OK? Let's try a practice word, MAN.

If the child produced a single word response to MAN, he was praised, and the experimenter stated several other words that he might have said. That is, for example, if the child's response was DADDY, the experimenter would say (approximately): "Good. That was a good one. Or you could have said other words too, like WOMAN, or BOY, couldn't you?" The test was then begun.

If the child produced no response, or a phrase, or if he repeated the practice word MAN, the instructions were reiterated and possible responses to the word MAN were given, followed by as many as three additional practice words.

The above procedure was used in the first session and, with slight modifications, in the second session as well.

Results and Discussion

After all data had been collected, a set of criteria was developed for the elimination of unacceptable responses. Such responses included any of the following: omissions, repetitions of the stimulus word, use of the stimulus word in a phrase response, use of items given as examples, repetition of a response, idiosyncratic name responses (e.g., "Dr. Jones"), responding with a letter, repetition of immediately preceding stimulus word, strong associate of another stimulus or response word, nonsense rhymes, unintelligible responses, and "room cue" responses (e.g., the child hesitates, gazes about the room, focuses on the window, and says "window"). Approximately 15 per cent of all responses were thus judged to be unacceptable,

with nearly half of these being response omissions. No age or sex difference in this respect was noticeable. There was considerable variation from child to child in the number or responses judged unacceptable, with a few subjects giving nearly all unacceptable responses. However, no subject was excluded from the sample unless all his responses were unacceptable or unless the session had to be terminated because of a lack of cooperation.

Table 4, which begins on the following page, shows response frequencies for all acceptable responses given to each word by three or more subjects broken down by age and sex. Hopefully, data will be compiled in the near future for more subjects of these ages and for additional words. The procedures employed here appear reasonably effective for this purpose. In the meantime, the data shown in Table 4 offers some help for investigators concerned with verbal learning and behavior of preschoolers.

Table 4*

Frequency of Associative Responses by Preschoolers to 44 Common Words

	3-year-olds		4-year-olds		
	Males (N=24)	Females (N=24)	Males (N=24)	Females (N=24)	Total (N=96)
BED					
sleep	9	12	9	10	40
bedtime	2	0	0	1	3
BIRD					
fly	10	7	7	11	35
tree	0	2	2	1	5
sing	0	0	2	1	3
tweet	0	3	0	0	3
wing	2	0	0	1	3
BLACK					
white	0	0	5	1	6
dark	0	2	3	0	5
color	2	1	0	2	5
red	0	1	1	2	4
paint	1	0	2	1	4
brown	1	0	1	2	4
crayon	1	2	0	0	3
outside	0	2	0	1	3
blue	1	0	0	2	3
BOY					
girl	1	2	4	7	14
walk	3	1	2	1	7
BREAD					
eat	10	12	10	8	40
butter	1	0	3	1	5
sandwich	0	1	1	2	4
toast	2	0	0	1	3

*The stimulus words are shown in upper case on the left with responses listed immediately below them. Only responses given by a total of three or more subjects to a particular word are included.

Table 4 (cont.)

	3-year-olds		4-year-olds		
	Males (N=24)	Females (N=24)	Males (N=24)	Females (N=24)	Total (N=96)
CAR					
drive	8	6	7	8	29
ride	1	0	1	2	4
truck	1	1	1	1	4
wheel	2	1	0	1	4
people	0	2	0	1	3
CARRY					
baby	1	1	6	1	9
heavy	0	1	3	1	4
people	0	1	1	1	3
food	2	0	0	1	3
walk	1	1	1	0	3
bag	1	1	1	0	3
CHAIR					
sit	6	9	7	11	33
sit down	5	5	4	3	17
table	1	2	3	1	7
COLD					
warm	3	1	3	3	10
freeze	3	3	1	3	10
jacket	3	2	1	2	8
outside	0	4	3	0	7
hot	3	1	1	1	6
winter	1	2	1	1	5
snow	0	0	1	2	3
coat	1	0	1	1	3
COME					
here	3	2	3	2	10
home	2	2	2	1	7
house	0	0	2	1	3
CRY					
sad	5	3	4	6	18
happy	1	1	2	3	7
tear	1	2	1	1	5
baby	0	1	1	2	4
spank	2	1	1	0	4
eyes	2	2	0	0	4

Table 4 (cont.)

	3-year-olds Males (N-24)	3-year-olds Females (N=24)	4-year-olds Males (N=24)	4-year-olds Females (N=24)	Total (N-96)
DOCTOR					
nurse	2	3	6	4	15
sick	1	2	3	6	12
better	2	2	1	1	6
shot	0	4	1	0	5
people	1	2	0	0	3
hospital	0	0	0	3	3
work	0	0	2	1	3
fix	2	1	0	0	3
DOOR					
open	6	5	8	3	23
close	3	4	1	3	11
handle	1	0	2	2	5
shut	1	0	0	3	4
outside	1	0	0	2	3
wall	0	0	3	0	3
FLOWER					
smell	3	3	2	4	12
grow	3	0	4	2	9
pretty	1	5	2	1	9
pick	1	0	1	2	4
outside	2	2	0	0	4
FRUIT					
eat	7	6	8	4	25
apple	1	1	2	2	6
orange	1	0	0	1	
GALLOP					
horse	10	5	10	10	35
skip	0	0	1	2	3
HAMMER					
nail	8	3	5	4	20
saw	1	1	0	3	5
wood	1	2	0	0	3
screw	1	1	1	0	3
screwdriver	2	0	1	0	3

Table 4 (cont.)

	3-year-olds Males (N-24)	3-year-olds Females (N=24)	4-year-olds Males (N=24)	4-year-olds Females (N=24)	Total (N=96)
HAND					
finger	3	2	2	3	10
arm	3	1	0	3	7
touch	1	0	2	2	5
shake	2	1	1	0	4
train	0	3	0	0	3
pick	0	0	3	0	3
table	1	2	0	0	3
HAPPY					
sad	2	3	6	4	15
smile	1	4	4	5	14
cry	0	2	0	1	3
mad	1	0	0	2	3
HEAVY					
carry	3	0	4	6	13
box	0	1	5	0	6
light	1	0	1	1	3
down	2	1	0	0	3
clothes	0	0	2	0	3
LAMP					
light	1	1	9	5	16
turn on	2	2	4	2	10
lamb	0	1	0	2	3
LION					
tiger	4	2	5	5	16
roar	4	4	1	3	12
bite	3	4	2	1	10
zoo	3	0	2	1	6
scare	0	1	1	2	4
elephant	4	0	0	0	4
bark	1	1	1	0	3
eat	0	1	2	0	3
growl	0	1	2	0	3
LISTEN					
talk	0	2	2	2	6
ear	1	1	2	1	5
quiet	1	1	2	0	4
sit	0	2	1	0	3
people	0	1	1	1	3

Table 4 (cont.)

	3-year-olds Males (N-24)	3-year-olds Females (N=24)	4-year-olds Males (N=24)	4-year-olds Females (N=24)	Total (N=96)
LOUD					
quiet	4	2	4	3	13
scream	3	2	3	2	10
talk	1	3	0	3	7
hear	2	0	1	1	4
soft	1	0	2	1	4
yell	0	1	1	2	4
bird	1	2	0	0	3
MOON					
dark	4	4	4	3	15
star	2	3	2	3	10
sky	3	0	4	2	9
night	1	0	1	6	8
sun	1	2	1	0	4
sleep	0	2	0	1	3
MOUNTAIN					
climb	4	1	4	5	14
big	1	1	2	1	5
water	2	3	0	0	5
cat	1	1	0	1	3
car	2	0	0	1	3
fountain	2	0	1	0	3
hill	1	0	2	0	3
MUSIC					
record	3	1	2	1	7
piano	1	1	3	2	7
hear	2	1	1	2	6
dance	2	1	1	1	5
guitar	2	1	1	1	5
record player	1	1	1	2	4
sing	1	1	1	1	4
turn on	1	1	2	0	4
listen	1	1	1	0	3
NEEDLE					
sew	3	3	3	7	16
thread	1	1	2	2	6
stick	2	1	2	1	4
hurt	3	0	1	0	4
knit	0	2	0	1	3
hammer	1	0	0	2	3

Table 4 (cont.)

	3-year-olds Males (N=24)	3-year-olds Females (N=24)	4-year-olds Males (N=24)	4-year-olds Females (N=24)	Total (N=96)
OCEAN					
water	4	3	2	3	12
swim	3	4	2	2	11
boat	5	0	2	2	9
lake	2	0	2	2	6
sea	0	0	4	1	5
PRETTY					
dress	2	4	2	1	9
girl	0	2	0	2	4
star	3	0	1	0	4
flower	1	1	1	1	4
QUIET					
sleep	4	6	3	2	15
shhh . . .	3	2	1	2	8
bed	3	2	0	2	7
loud	0	1	2	2	5
down	0	1	1	2	4
talk	0	0	2	1	3
RIVER					
water	4	3	4	4	15
boat	4	2	1	2	9
swim	1	1	1	4	7
fish	1	1	3	0	5
car	1	1	1	1	4
lake	2	0	0	2	4
RUN					
fast	5	3	5	3	16
walk	2	0	0	3	5
slow	1	1	1	2	5
stop	1	0	0	2	3
play	1	1	0	1	3
outside	2	1	0	0	3
SAD					
cry	6	6	4	5	21
happy	5	2	5	3	15
mad	1	0	2	2	5
tear	0	1	0	2	3
SALT					
pepper	3	4	7	9	23
food	2	1	2	1	6
dinner	4	0	1	1	6
eat	0	1	1	1	3

Table 4 (cont.)

	3-year-olds		4-year-olds		
	Males (N=24)	Females (N=24)	Males (N=24)	Females (N=24)	Total (N=96)
SHEEP					
"baa"	5	7	1	1	14
lamb	3	2	4	4	13
dog	0	2	2	0	4
cow	2	0	1	1	4
bed	1	0	0	2	3
SIT					
down	8	9	7	8	32
chair	4	3	5	3	15
eat	1	0	2	0	3
SLEEP					
bed	4	2	7	2	15
wake	3	2	2	4	11
close your eyes	0	2	2	2	6
tired	1	3	0	1	5
SLOW					
fast	2	4	3	5	14
down	3	6	3	1	13
walk	5	1	3	2	11
car	2	1	1	1	5
drive	1	1	1	1	4
stop	1	0	1	1	3
TABLE					
eat	9	9	10	10	38
chair	0	3	3	3	9
sit	2	1	2	2	7
floor	3	0	0	0	3
TELL					
telephone	1	0	2	3	6
show	0	0	3	1	4
car	3	0	0	0	3
something	2	0	0	1	3
T.V.	1	0	0	2	3
talk	0	0	1	2	3
THIRSTY					
drink	4	8	5	6	23
drink of water	3	0	2	5	10
water	1	1	0	4	6
hungry	2	1	3	0	6
glass	1	1	1	0	3
not thirsty	0	0	1	2	3

Table 4 (cont.)

	3-year-olds Males (N=24)	3-year-olds Females (N=24)	4-year-olds Males (N=24)	4-year-olds Females (N=24)	Total (N=96)
UP					
down	7	4	8	9	28
sky	1	0	1	3	5
hill	2	1	1	0	4
elevator	1	1	1	0	3
WILD					
animal	1	1	1	1	4
water	1	0	1	2	4
fight	0	0	2	1	3
dog	0	2	0	1	3

Experiments VII and VIII

Modifying Preschoolers' Use of Conceptual Encoding Processes: Two Learning-to-Learn Experiments

The central concern of Experiments VII and VIII was the modifiability of operations by which three- and four-year-old children encode and retrieve categorizable information. Recall from Experiment IV of this report that free recall by kindergarten and first-grade children was facilitated when the to-be-remembered items were categorizable, whereas with preschoolers performance on categorizable lists was no better than for lists of unrelated words. Although the precise mechanisms by which such facilitation occurred for the older children was not explicated, it was assumed that they were operating during the encoding of the items, at the time of retrieval, or both. Specifically, the working hypothesis adopted here was that the older children were more likely than were the younger ones to produce and rehearse category labels which, during testing, were recalled and used to assist in retrieving the appropriate instances. One apparently plausible explanation of this age difference is that the operations by which the older children capitalize on the categorizability of items constitutes a strategy acquired through experience, and especially through experience provided in kindergarten and first-grade classes. If so, then one would expect that such a strategy could be acquired earlier (i.e., by preschoolers) provided that the relevant experiences were provided them. This notion was tested in Experiments VII and VIII by designing and implementing what might be called a "conceptual encoding curriculum" for preschoolers.

The two experiments, which were highly similar to one another, involved the random assignment of preschool children to experimental and control conditions. In each experiment the experimental children were exposed to a series of training exercises over a period of several months and more than 20 separate sessions. The training exercises were cast in the form of games, and were designed to increase the extent to which the children became aware of conceptual relationships among items and produced category labels to individual instances during encoding and retrieval. At the conclusion of training a series of "tests" were administered to examine the effects of training. The test tasks were designed with particular attention to effects that generalized beyond the particular tasks and materials used during training, i.e., to assess learning-to-learn effects.

Method: Experiment VII

Subjects

Forty-six children enrolled in a private nursery school-day care center in Evanston were randomly assigned to two groups, each containing

23 Ss. One S was lost from each group prior to the end of training, reducing the groups to 22 Ss each. The experimental group consisted of eight males and 14 females with a mean CA of 47 months and a range of 39-56 months at the commencement of training. The control group consisted of 10 males and 12 females with a mean CA of 48 months and a range of 36-59 months. The limited socio-economic data available indicates a wide range socioeconomically, with major representation from the lower-middle and upper-lower classes.

Training Exercises and Procedure

Seven different training exercises (or types of exercises) were employed, each involving a content that consisted of instances of several common categories. The particular categories employed during training were selected randomly from a pool of 20 categories that were judged as being relatively familiar ones to most preschoolers. The 10 training categories, with the number of instances for each were as follows: birds (6), bodies of water (3), bugs (7), colors (4), feelings (7), furniture (6), parts of buildings (8), parts of the body (6), things to eat with (6), and vegetables (5). The remaining categories (new categories) were: appliances, beverages, clothing, dwellings, fruit, mammals, money, musical instruments, occupations, and means of transportation.

Exercise A: Category naming and grouping -- pictorial. This was the initial exercise used. Pictures were introduced and the S was asked to provide the picture label as well as the appropriate category name. After the introduction of several instances of each of several categories, the S was asked to classify instances by category. Labels for instances and categories were supplied as necessary. Three sessions were occupied by this exercise, the sessions differing only in the particular categories and instances that were used.

Exercise B: The puppet show. This exercise, unlike the others, was not in a game format. Rather, the experimenter-teacher depicted a variety of feelings (anger, surprise, etc.) by putting on an amateur puppet show. Periodically the children were asked to identify the feelings that were appropriate under various conditions that were depicted. Each child participated in one session.

Exercise C: Remembering categories and instances -- verbal. Instances of various categories were spoken, followed by questions as to whether or not particular categories had been represented. For instance, the experimenter might say ANT, BUTTERFLY, HAPPY, and PROUD, followed by queries as to whether any names of BUGS or BIRDS or FEELINGS had been said, then by questions as to which particular instances had been presented. Feedback was given following each response. Five sessions of this exercise were used, with specific content varying from one session to the next.

Exercise D: Organizing pictures by category. In this game, one experimenter displayed pictorially several instances of each of several categories, and instructed the S to look at and think about the pictures so that he could remember them. A second experimenter worked along with the S, suggesting that the pictures be organized by category, modeling such organization at the beginning of the session, and then encouraging the S to take the initiative in further organization. Finally, the S was tested for recognition of categories and recall of instances, with feedback and occasional coaching. Three such sessions per S were used.

Exercise E: Recognition of category membership. This game consisted of the oral presentation of sets of three or four words, one of which did not belong in the same category as the rest. The Ss were asked to identify the one not belonging. Each S participated in three such sessions.

Exercise F: The store games. A "play store" was stocked with items (pictures of items) from a variety of categories. The S was told of two categories of items stocked by the store, and sent to by items from these categories and to find out what other categories of items were stocked. Subsequent coaching and quizzing emphasized memory of instances together with the categories into which the various items fell. Six sessions per S were given, with variations both in content and in specific activities involving the store.

Material and procedural details. Small group sessions were used for the Puppet Show and Recognition of Category Membership, whereas the remaining exercises were individual ones. Scripts were prepared for each session and the various experimenters met frequently to ensure uniformity of procedures and adherence to the exercise formats. The scripts were designed so that the various categories and instances occurred with roughly equal frequency during the course of training. All Ss received the exercises in the same predetermined order, beginning with a session of Exercise A above and ending with sessions of Exercise F. The exercises were conducted by two of the regular nursery school teachers together with three Northwestern University students -- all females.

Most of the pictures used during training were in color and were either taken from various published educational game materials or were cut out of magazines and mounted on stiff paper. No attempt was made to achieve uniformity in size or style of pictures.

All training was conducted during school hours in a schoolroom reserved for that purpose. Training sessions generally lasted about 15 minutes and were distributed throughout a period of about two months. To increase motivation for participation, the experimental Ss were given frequent praise and a chart was kept on which a star

was displayed for each session completed by each S. During the periods of training the control Ss participated in the regular activities of the nursery school. Although no attempt was made to limit the school experiences of the control Ss, it appeared that there were few activities for them that seemed likely to influence the process on which the training exercised focused. Note also that the control Ss were quite familiar with the experimenter-teachers, having interacted with them prior to commencement of the experiment and continuing such interaction in a variety of non-training situations during the experimental period.

Test Tasks and Procedure

Because the school year was nearly over by the end of the training phase, time remained for only three testing sessions, making it impossible to investigate all potential training effects of interest. Since major interest was in relatively generalized learning-to-learn effects concerning memory-encoding processes, the test tasks reflected that interest at the expense of more specific training effects. The tasks used are described below.

Free study-recall of pictured objects. This is a task modeled after one used by Moely, et al. The Ss were first asked to identify pictures of the following objects: cat, shoe, train, strawberries, car, horse, apple, pants, banana, hat, elephant, and airplane, with the experimenter supplying the correct name as necessary. (Note that these objects fall into four main categories, none of which were employed during training: animal (or mammal), clothing, means of transportation, and fruit. The S then was given the pictures, arranged in the above order, and told to put them on the table in front of him and look at them. He was told also that he could move the pictures around any way he wished to help him remember them, and that the experimenter would do some work at the other side of the room and later return to find out how many pictures the S could remember without looking. The experimenter sat down across the room and pretended to work while surruptitiously abserving the S for signs of picture rearrangement and overt rehearsal. After a two-minute study period the experimenter returned and, after noting the picture arrangement, collected the pictures and asked the S to tell as many of the pictures as he could. It was predicted that experimental Ss would show more picture rearrangement by categories and better recall.

Paired associate learning (PAL) -- facilitation. In this task the S was given one initial presentation followed by two anticipation trials on each of two six-pair lists. In each list each pair consisted of two common instances of the same category. In List 2 the categories and instances were among those used during training, whereas in List 1 only instances of new categories were used. List 1 was given first to

Table 5

PAL -- Facilitation Word List Used in Experiment VII

List 1			List 2		
PENNY	-	DOLLAR	ROBIN	-	EAGLE
SWEATER	-	SOCKS	SURPRISED	-	HAPPY
TEACHER	-	DOCTOR	RIVER	-	LAKE
COW	-	DOG	GREEN	-	YELLOW
GUITAR	-	PIANO	ANT	-	SPIDER
PINEAPPLE	-	ORANGE	SPOON	-	CUP

all Ss. The lists are shown in Table 5 . It was anticipated that the performance of experimental Ss would exceed that of the controls on both lists, and that this difference would be greatest on List 2, where both specific and general training effects would operate. This task was administered during Session 1, closely following the Free Study-Recall Task.

Free recall of conceptually related words. A list (List 1) of 14 common words each was presented orally at a four- to five-second rate followed by a single unpaced recall test. After one minute that procedure was then repeated with a second 14-item list (List 2). These lists are shown in Table 6 . Within each list the first eight items consisted of two instances of each of four categories, and the next four words consisted of the relevant category labels. The last two words in each list were filler items included to reduce the recency effect for the first 12 "critical" items. List 2 included categories and instances used in training, whereas List 1 was made up of new items.

Between presentation of items and recall testing the Ss were given a 10-second period during which they were told that they "can think about those words for a little while."

If the training had the anticipated effects on the use of conceptual processes, then experimental Ss should exceed controls on both lists, with a greater difference for List 2 since test items were identical to training items. No other task was administered in Session 2.

Recognition memory (RM) for pictures. Twelve pictures of common objects were presented for four to five seconds each, followed by the oral presentation of a 20-item test list. The to-be-remembered objects were: BOAT, TRAIN, BUS, TRUCK, AIRPLANE, BICYCLE, COAT, BELT, SHOE, HAT, DRESS, and SWEATER, presented in that order for every S. The test list contained those same 12 items (verbally, rather than pictorially) plus six new unrelated words, ordered randomly. In addition, two new related items (PANTS and CAR) were included as the final test list items. It was reasoned that the expected training effects would result in better performance on the first 18 test items by experimental Ss, but poorer performance (more false recognitions) on the final two items.

PAL-interference. In Session 3, following the RM task, Ss were presented two PA lists using the same procedure as in the PAL facilitation task described earlier. List 1 contained the pairs NICKEL-FRUIT, BLOUSE-ANIMAL, ZEBRA-MONEY, and PEAR-CLOTHING, in which instances and categories (new categories) were inappropriately paired. Similarly, List 2 contained the pairs EAGLE-FURNITURE, GREEN-BIRD, TABLE-VEGETABLE, and CARROT-COLOR, wherein all instances and category

Table 6

Conceptually Related Word Lists Used
in Free Recall Task in Experiment VII

List 1	List 2
QUARTER	OWL
APPLE	BLUE
GIRAFFE	SAD
SHIRT	CORN
DIME	PIGEON
PEACH	RED
BEAR	SCARED
JACKET	BEANS
MONEY	BIRD
FRUIT	COLOR
ANIMAL	FEELINGS
CLOTHING	VEGETABLES
TREE	STORY
KING	KITE

names were among those used during training. To the extent that training had increased the general tendency of experimental Ss to implicitly produce category labels during encoding, these Ss were expected to encounter more interference, and thus perform more poorly, than controls. That difference would be greater on List 2 where highly specific interference was expected as well.

Material and procedural details. Wherever pictures were used in testing, they were different ones than had been used during training. Each child was tested individually by an experimenter who had not been involved in his training, and, with only a few exceptions, the tester was unaware of the S's group membership (experimental vs. control). Testing was carried out mainly in the room used for training. However, the control children also were familiar with this room and with tests of the sort used here by their prior participation in verbal learning experiments.

Results and Discussion: Experiment VII

Table 7 provides a summary of performance on the various test tasks. The data concerning rearrangement of pictures by category during the free study-recall tasks are not shown there because no such rearrangement occurred by an S in either group. The data shown in Table 7 do not suggest any dramatic effects of the training. Of the nine comparisons in that table, six differences are in the predicted direction, with one of these (PAL interference with old material) highly significant, t (42) = 2.87, $p < .01$. However, one of the differences in the opposite direction (FR of new words) also was significant, t (42) = 2.00, $p < .05$.

Clearly there is no evidence of any training effect that generalized beyond those materials employed during training. The one significant difference favoring the experimental Ss occurred for categories and instances that were used in training. Moreover, although in that case the task (PAL) had not been used in precisely the same way during training, certain of the training exercises involved highly similar activities. That is, Ss frequently were given an instance and asked to produce the appropriate category name, precisely what would have resulted in the interference in the test task in question. Thus, there is some evidence for a relatively specific training effect -- specific to the particular training content and highly related to training activity. Even this evidence is equivocal, however, since no significant differences were found for the remaining two tasks in which old categories and instances were involved. Can we be sure, then, that anything was learned during training? Did the Ss improve their performance on the training exercises themselves from session to session? Unfortunately, partly because of practical limitations, and because of the author's conviction that more general effects would be demonstrated, the evidence on this point is only anecdotal. The experimenter-teachers

Table 7

Summary of Performance on Test Tasks
in Experiment VII

Test Tasks	New vs Old Category	Experi-mental (E)	Control (C)	Direc. of Diff. Expected	Obtained
Free study-recall	New	6.32	6.09	E > C	E > C
FR	New	3.68	4.86	E > C	E < C*
FR	Old	3.55	3.64	E > C	E < C
PAL facilitation (Trials 1 & 2 combined)	New	8.32	7.55	E > C	E > C
PAL facilitation (Trials 1 & 2 combined)	Old	9.36	9.14	E > C	E > C
PAL interference (Trials 1 & 2 combined)	New	2.14	1.77	E < C	E > C
PAL interference (Trials 1 & 2 combined)	Old	1.18	2.27	E < C	E < C**
RM (Old words - new unrelated words)	New	9.64	9.05	E > C	E > C
RM (New related words - new unrelated words)	New	.04	-.05	E > C	E > C

involved were convinced that learning did occur. They reported that, with increased experience, the Ss became better able to classify various instances by category and to produce the appropriate category names.

The hypothesis that preschoolers can be taught to generally take greater advantage of conceptual relationships for encoding and retrieving information was not confirmed. Two broad conclusions are possible. The first of these is that the encoding and retrieval operations involved are not modifiable by experience for most three- and four-year-old children. The second possibility is that the hypothesis was not adequately tested in the present experiment. To evaluate the latter alternative, consider the strengths and weaknesses of the experiment. Were the training exercises inappropriate for their purpose? That seems unlikely. On the face of it, at least, the emphasis in these exercises was highly consistent with their purpose. To review, these exercises gave verbal instructions and practice in naming pictured objects, clustering objects on the basis of category, and producing category names for instances presented orally or pictorially. Perhaps the exercises were well-conceived but training was too brief. This seems unlikely. Even if one were to argue that 21 sessions were not sufficient to produce substantial effects, the effects, although perhaps relatively small, should at least be detectable. But perhaps the exercises were theoretically sufficient, but, in practice, the Ss did not attend adequately to them, thus the actual amount of training received was far short of that planned. If so, then one would expect that the experimenter-teachers would report such inattention and that Ss would resist participation, especially later during the training period. Neither was true. Although some inattention was reported, it was minimal. Moreover, the Ss typically were willing, and often were anxious to participate. Perhaps the desired effects were produced, but were not detected by the dependent measures. This again appears unlikely, since several such measures were used, and these tasks were modeled after ones used successfully in prior research. However, before reaching a firm conclusion regarding the modifiability of the processes in question it was decided that the experiment should be replicated, with certain alterations both in the training and the test phase. That was the purpose of Experiment VIII.

Experiment VIII: Method

Subjects

The Ss were 56 nursery school and day care children again selected from those attending the Tom Thumb Nursery School in Evanston, and randomly assigned to the experimental and control groups, with 28 Ss in each at the outset of training. The experimental group consisted of 14 males and 14 females, with a mean CA of 46 months, and a range of 36-57 months. The control group included 12 males and 16 females, with a mean CA of 46 months and a range of 37-56

months. Nine experimental and seven control Ss left the school at some point during either the training or the testing period, resulting in slightly differing numbers of Ss for the various tests used.

Training Exercises and Procedure

The exercises (A-F) used in Experiment VII were used again in Experiment VIII, with the same numbers of sessions for each of the exercises except for Exercise F (the store game) where only four (rather than six) sessions were employed in Experiment VIII. Three new exercises also were used. Exercise G was a small group game, with three sessions per S in which pictures of objects were distributed to the Ss, and each S was asked to decide which other S's picture his own picture went with, to name each picture, and to name the appropriate category. In Exercise H the Ss (individually) were told to listen for instances of particular categories in a short story to be read by the experimenter-teacher. After the story was read, the S was asked to recall all such instances that appeared. For instances not recalled a RM procedure then was used. One such session was used.

Exercise I involved the "play store" as in Exercise F, but the game in this case was one of planning and stocking the store to get ready for business. From an unorganized stack of pictures of items to be sold, the S was required, in each of three sessions, to organize items by category within the store, and to produce the names of the various items and categories.

Thus, during the five-month training period each S participated in 26 sessions involving nine different games. The content of the exercises was determined by randomly selecting nine categories from a pool of 18 categories with which the Ss were relatively familiar. The training categories were: BIRDS, BUGS, FEELINGS, FURNITURE, OCEAN ANIMALS, PARTS OF THE BODY, THINGS TO EAT WITH, TOOLS, and VEGETABLES. The remaining (new) categories included: ANIMALS, CLOTHING, FRUIT, MONEY, MEANS OF TRANSPORTATION, MUSICAL INSTRUMENTS, OCCUPATIONS, and PARTS OF A HOUSE.

The two nursery school teachers who served as experimenter-teachers in Experiment VII again served this function, assisted by two Northwestern undergraduate females. In other respects, the materials and methods for the training were the same for the two experiments.

Test Tasks and Procedures

Testing was carried out in eight individual sessions per S, over a period of about 10 weeks. As in Experiment VII, the tests

were designed to examine generalized effects of training on operations by which the Ss encoded and retrieved categorizable information. However, a wider variety of possible effects were examined in this study, and greater attention was given to relatively specific effects. Except as noted below, the details regarding materials and procedures were as in Experiment VII.

Producing category labels. This was one of only two test tasks administered prior to the completion of the training exercise, at a time when the Ss had completed about three-fourths of the training sessions. It also was one of only two tests that were administered by individuals who had worked with the experimental Ss during training.

The following list of words was read, in this order, to each S at about a five-second rate: SPIDER, TABLE, BED, FLY, HAND, CHAIR, OWL, SAD, BEE, DESK, ANT, and SCREWDRIVER. Note that the list contains instances of six of the training categories, beginning with the categories BUG and FURNITURE. Each S was instructed to say "bug" to every such instance, and "furniture" to every such instance. There were four instances of each of those categories. No mention was made of the categories into which the remaining four words fell. Our interest was in the effects of training on Ss' ability to produce the requested category labels (a very specific effect) and on the tendency to produce appropriate category labels even when not specifically requested to do so (a somewhat more general effect).

Associative responses: I. This is the second of the two test tasks administered prior to the end of the training period and by testers who had participated in the training. Each S was read aloud a series of 18 words of which nine were category instances employed during training and nine were instances of the nine new categories. Half the Ss in each group, randomly selected, received the old items first and half received them last. The nine old items were HAMMER, NOSE, SPOON, CORN, SPIDER, HAPPY, CHAIR, OWL, and FISH. The new items were DOOR, HORSE, BANANA, PENNY, MILK, FIREMAN, AIRPLANE, GUITAR, and SHIRT. All Ss were told, "We're going to play a word game together. I'm going to think of a word. When I tell you my word, I want you to think of a word that goes together with my word." Although rate of presentation was not precisely controlled, if no response was given for about five seconds, the tester read the word again. If there still was no response, the tester went on to the next word. This test procedure was similar to, but not identical with, certain of the training exercise components wherein Ss were presented a picture and asked to identify other pictures that "go together" with it. If the experimental Ss produced more category labels as responses to the old items than did control Ss, this would indicate

learning, but of a fairly specific sort. However, if that difference also prevailed for the new items, it would indicate a more general training effect, i.e., one that generalized across categories.

Associative responses: II. The task and procedures here were identical with that just described, except that the items selected from the old (training) categories were ones not used in training. This task was administered during the sixth testing session, several weeks after the training period had ended. The task allowed for the examination of training effects that generalized to new instances within old categories as well as to new categories.

PA learning and retention -- interference. In this task the Ss were given PA training with pairs in which old instances served as stimulus items and were paired with old but inappropriate category labels as follows: OCTOPUS - BIRD, FOOT - TOOL, CHAIR - FEELINGS, and PEAS - BUG. The training procedure consisted of sets of trials in which a drop-out procedure was used. That is, only items missed on previous trials were presented, until all items were answered correctly, and the tester moved to another set of trials using the same procedure. All Ss were run to a criterion of two consecutive correct trials or until 10 sets of trials had been completed, whichever came first. All presentation was oral, and began with an initial trial in which the pairs were presented as follows: "When I say _____, you say _____." On later trials the tester said, "What word goes with _____?" If the correct response was given, the S was so informed. If no response or an incorrect response was given, the S was told the correct response. The order of presentation varied randomly between sets.

A retention test was given one week (in most cases) after learning, and again five-seven weeks after learning. In both cases the procedure simply involved reminding the S of the earlier learning session, then presenting each stimulus item orally at about a five-seven second rate, with no feedback.

The rationale for the task was derived from an interference theory of forgetting in which proactive interference is seen as an important mechanism. If, as anticipated, the training resulted in the strengthening of associations between instances and appropriate category labels, then PAL involving inappropriate labels should be more difficult. Thus, the experimental Ss would be expected to require more trials to criterion. Further, once all Ss reach criterion (an equal level of original learning), forgetting ought to be more rapid for the experimental Ss because of the recovery of the highly practiced appropriate category names.

PAL -- facilitation. The Ss were given five PAL trials with an eight-pair list using the same procedures used in the PAL-facilitation task in Experiment VII. The pairs consisted of old instances paired with old category labels (or the key words in the labels) as follows: HAMMER-TOOL, PARROT-BIRD, CORN-VEGETABLE, FOOT-BODY, TABLE-FURNITURE, ANT-BUG, PROUD-FEELING, and SHARK-OCEAN. In this case superior performance by the experimental Ss presumably would be expected if training had increased the strength of the associations between the instances and their category labels whether or not any more general tendency to produce category labels had been affected.

FR of categorizable words. All Ss were given four trials with each of two nine-item lists, a new list and an old list. The old list contained three old instances of each of three old categories, as follows: OWL, ROBIN, EAGLE, HAPPY, SAD, SCARED, SPOON, KNIFE, and FORK. The new list consisted of three instances of each of three new categories: TEACHER, DOCTOR, MAILMAN, MILK, WATER, COFFEE, DOLLAR, NICKEL, and PENNY. Instances of a category were always blocked for presentation, but the order of presentation of blocks and of instances within blocks varied randomly across trials. The items were read to the Ss at about a three-second rate. Half the Ss received the old list and half the new list first, and the lists were administered during different testing sessions (sessions 2 and 3).

Recall that in Experiment IV, described earlier in this report, the FR of preschoolers, unlike school children, for a blocked-categorizable list was no better than for an uncategorizable list. Would the training given the experimental Ss result in an increased tendency for those Ss to capitalize on such relationships? If so, would that effect be confined to the concepts included in the training exercises or would it generalize to new material? These are the questions that the FR test task was designed to answer.

Free study-recall of categorizable pictures. This task, administered in session 5, was highly similar to the free study-recall task of Experiment VII. Pictures (mounted on 5" x 7" cards) of three instances of each of four new categories were arranged in a stack in the following order: PIG, GUITAR, CAR, HAT, CAT, TRAIN, PIANO, SHOE, AIRPLANE, DRUM, HORSE, and DRESS. The S was seated at a table on which 12 5" x 7" spaces were marked in an array containing four columns and three rows, with distinct separations between the columns. This arrangement was designed to easily accomodate an arrangement of the pictures by category. Each S was told that he was to place the cards in the spaces provided, and this placement was illustrated by placing the first four of the pictures along the top row of the array of spaces. The S was then given the remaining

eight pictures and asked to place them. He was told that he could place them and/or rearrange them in whatever way would help him remember them, and that after a while the experimenter would return, collect the pictures, and then find out how many could be remembered. The tester then retired to the far side of the room for three minutes. After one minute of this study period the tester said, "Remember, you can pick the pictures up and put them in different spaces if you think it will help you to remember the pictures better." During the study period the tester recorded verbal rehearsal and picture arrangement, then returned, noted the final picture arrangement, collected the pictures, and tested the S's recall.

Picture categorization. Two sets of 12 pictures were used in this test task, administered in session six. The set used first with all Ss consisted of four instances of each of three new categories, namely, DOG, SWEATER, APPLE, BANANA, LEMON, SQUIRREL, JACKET, PANTS, COW, ELEPHANT, SOCK, and ORANGE. The first three of the pictures listed above were placed on a table before the S. Then each of the remaining pictures, in the randomly determined order shown above, was shown to the Ss, with instructions to tell which of the three "standards" it "goes with best." After all pictures in the set had been placed in this fashion, the pictures were collected and the second set was presented in exactly the same fashion. The objects shown in the second set were BED, SAW, BEE, BEETLE, COUCH, HEDGER, BUTTERFLY, CHEST, CHAIR, GRASSHOPPER, HAMMER, and DRILL. Half were new and half were old instances of old categories. After completion of that set, the four members of each of the six categories (three new and three old) in turn were placed before the S who was asked first if they went together and then, after it had been established that they did, to supply a category label. For example, the pictures DOG, SQUIRREL, COW, and ELEPHANT were displayed first and the tester asked, "Do you think all these pictures go together? If the S answered affirmatively, the tester agreed, then said, "I want you to think of a word that tells us what they all are. What do we call things like these?" If the S had not agreed that they went together, the tester said, "Yes, these pictures do go together," then asked for the category label.

Category recognition -- verbal. In the final test task 12 sets of three words each were presented orally, with the S instructed to identify the member of each set that "does not belong with the other words." Six of the sets presented each consisted of two old instances of an old category and one distractor, and the remaining sets each included two instances of a new category along with a distractor. The order of sets and the order of words within sets were determined randomly and were the same for all Ss. These sets, in the order presented, were (1) BOAT, DIME, TRAIN: (2) FOOT, OWL, HAIR;

(3) COWBOY, GRAPES, STRAWBERRY; (4) CARROT, HAMMER, BEANS; (5) OCTOPUS, SHARK, HAPPY; (6) EYES, KNIFE, CUP; (7) HAT, BELT, FLOOR; (8) JUICE, WATER, DRUM; (9) LION, MITTEN, RABBIT; (10) BUTTERFLY, SPIDER, FORK; (11) GIRAFFE, PENNY, QUARTER, and (12) CORN, LAMP, BED.

Results and Discussion: Experiment VIII

Table 8 provides a summary of performance of the two groups on the test tasks. The reliability of the difference between means was examined by means of *t* tests. In five such cases the differences were significant, and all of these favored the experimental *Ss*. In fact, 19 of the 21 differences shown in Table 5 were in the predicted direction, although several were of very small magnitude and most did not approach significance.

The largest differences were found in producing category labels. Here the experimental *Ss* were far more likely to correctly apply the labels BUG and FURNITURE to the eight instances of these categories while avoiding the inappropriate use of these labels, $t\ (45) = 2.81$, $p < .01$. Also, they far more often produced correct category names for the remaining four instances, $t\ (45) = 3.49$, $p < .01$. In the first of the associative responding tasks the experimental *Ss* exceeded the controls in the number of appropriate category labels produced as associative responses, and that difference was highly significant in the case of old categories, $t\ (45) = 2.88$, $p < .01$. The second associative responding task yeilded a similar pattern of results, except that the significant difference occurred for the new categories rather than the old, $t\ (39) = 2.21$, $p < .05$. Note that this was the only comparison involving the new categories in which a significant difference was found. The failure of the difference for the old categories to reach significance is more understandable if one recalls that the instances used were new ones, i.e., had not been used during training.

The only learning task in which performance significantly favored the experimental group was the PAL facilitation task, and particularly, the first trial of that task, $t\ (43) = 2.18$, $p < .05$. This difference by itself could be accounted for on the basis of the strengthening during training of specific associations between instances and category labels, rather than by any more general tendency to produce category labels. The fact that the difference diminished over trials is uninterpretable because of a ceiling effect. No differences found for the PA interference tasks were significant, although all were in the direction predicted on the basis of the interference hypothesis as outlined earlier.

Table 8

Mean Performance on Test Tasks in Experiment VIII

Test Task	New vs Old Category	Experimental (E)	Control (C)	Direct. of Diff. Expected	Direct. of Diff. Obtained
Producing category labels					
Correct-incorrect use of designated labels	Old	4.48	1.50	E > C	E > C**
Correct use of nondesignated labels	Old	1.39	.29	E > C	E > C**
Associative responses: I					
Category label responses	Old	.57	.08	E > C	E > C**
Category label responses	New	.30	.08	E > C	E > C
Associative responses: II					
Category label responses	Old	2.15	1.43	E > C	E > C
Category label responses	New	1.95	.71	E > C	E > C*
PA interference					
No. correct on Trial 1	Old	.86	1.12	E < C	E < C
Errors -- all trials	Old	12.79	12.12	E > C	E > C
48-hr. retention	Old	2.60	2.67	E < C	E < C
5-7 wk. retention	Old	1.10	1.80	E < C	E < C
PAL -- facilitation					
No. correct on Trial 1	Old	5.13	3.56	E > C	E > C*
No. correct -- all trials	Old	28.23	27.96	E > C	E > C

Table 8 (continued)

Test Task	New vs Old Category	Experimental (E)	Control (C)	Direc. of Diff. Expected	Direc. of Diff. Obtained
FR of categorizable words					
No. correct on Trial 1	Old	3.68	3.26	E > C	E > C
No. correct -- all trials	Old	18.14	19.57	E > C	E < C
No. intrusions from appropriate categories	Old	1.91	.78	E > C	E > C*
No. correct on Trial 1	New	3.59	2.78	E > C	E > C
No. correct -- all trials	New	19.68	18.30	E > C	E > C
No. intrusions from appropriate categories	New	1.82	1.43	E > C	E > C
Free study-recall					
No. recalled correctly	New	5.24	5.04	E > C	E > C
Picture categorization					
No. correctly categorized	Old	8.67	8.62	E > C	E > C
No. correctly categorized	New	8.71	8.10	E > C	E > C
Category recognition -- verbal					
No. correctly recognized	Old	3.32	2.86	E > C	E > C
No. correctly recognized	New	3.11	3.29	E > C	E < C

The FR and free study-recall data are of particular interest in view of the results of Experiment IV, which provided much of the impetus for the training experiment. That is, the training exercises were designed especially to increase the extent to which preschoolers would capitalize on conceptual relationships among to-be-remembered items during encoding and retrieval. In that sense, then, these two test tasks were the most critical of those used, and no evidence for the expected effect was found, even when the instances and categories involved were identical to those used in training. However, one difference (not shown in Table 8) between the experimental and the control Ss on the FR task did emerge, and that concerned intrusions during recall. Nearly all intrusions were instances of the categories represented by the to-be-remembered items. In the case of the old categories, such intrusions were significantly more frequent among the experimental ($\bar{X}$ = 1.91) than for the control ($\bar{X}$ = .78) group, t (43) = 2.27, $p < .05$. For the list involving new categories the respective means were 1.72 and 1.43, a difference that did not approach significance. Thus, at least for the old items, it appeared that the training did increase conceptual encoding, but that this increase was not beneficial to the Ss' FR performance. In fact, it was detrimental in that the experimental Ss were similar to the control Ss in errors of omission, but higher in errors of commission (i.e., intrusions). Clearly, this was not the effect that was anticipated, and suggests the need for a closer examination of the processes by which older Ss ordinarily derive benefit from the existence of conceptual relationships among to-be-remembered items.

Table 8 shows no significant differences between groups in the final two tasks. However, in the case of the picture categorization task the lack of a difference may reflect a ceiling problem, since the task proved to be an extremely easy one. One significant difference, not shown in Table 8, did emerge for that task. When, at the close of the session, Ss were asked to produce category labels for the sets of conceptually related pictures, the experimental Ss produced more of the particular labels used during training for old category instances than did the control Ss. Such a highly specific training effect is of little theoretical interest, but it does indicate that learning of specific labels did occur during training.

General Discussion: Experiments VII and VIII

The data from Experiment VII did not suggest a generalized training effect on preschoolers' encoding and retrieval operations. In fact, that experiment did not provide convincing evidence of any training effect, in that a relatively restricted range of test tasks was employed, and the pattern of results of them was somewhat inconsistent. Since the results of Experiment VIII were far more

informative, and since the training program was more intensive, the following discussion will center most heavily on that experiment.

Experiment VIII provided clear evidence of training effects. The first question to which this discussion will be addressed concerns the specific nature of these effects. Can the effects be accounted for by assuming an increase in the strength of specific associations during training or is there evidence for the acquisition of more general response tendencies or strategies? Of the eight significant differences favoring the experimental Ss (one in Experiment VII and seven in Experiment VIII), all but one can be explained on the basis of specific associations, in that they occurred exclusively with instances and categories that had been employed during training. More specifically, all could have resulted from specific training in producing those particular category labels appropriate to those specific instances. The single exception to this occurred in the case of the second associative responding test task in Experiment VIII, where experimental Ss responded more frequently with category labels even when the presented words were instances of categories not included during training. It appears, therefore, that the primary training effects were those that would be expected to occur in the typical PAL task. That is, stimuli (instances) were differentiated, response terms (category labels) were acquired and/or strengthened, and the associations between the instances and categories were strengthened. As a result, in the various test tasks involving old items, the experimental Ss more often produced those category labels when instances were presented. There is some evidence, although not as convincing, that beyond this specific effect the experimental Ss acquired to some degree the concept of a concept or category as well as the tendency to produce category labels to items not encountered during training. That more interesting effect seems, at best, to have been a relatively weak one. In short, the data suggest that such training of preschoolers as was employed here does substantially increase conceptual encoding when the material to be learned is identical to that used during training, but only slightly, if at all, for other material.

Now consider the consequences of the above effects on performance in learning or memory tasks. For PAL, using old instances as stimuli and old category labels as response terms, the consequences were beneficial, just as would be expected. More interesting is the case of FR. Here the consequences were detrimental, in that intrusions from the appropriate categories were more frequent for experimental Ss without a corresponding increase in correct recall. Although the data do not explicate the reasons for this unexpected result, the following speculative account is offered. When older children and adults are engaged in learning a categorizable list,

they produce category labels and rehearse the presented instances. The former is useful as a retrieval cue, and the latter provides a basis for distinguishing between retrieved instances that had and those that had not been presented. Thus, both are necessary to derive full benefit from the conceptual relationships present on the list. In the present experiment only one was emphasized, namely, the production of category labels. Thus, at recall the S did have the label available as a retrieval cue but did not have available a satisfactory basis for the required discrimination. That basis ordinarily may be the perceived situational frequency difference between presented and nonpresented items that builds up through simple rehearsal. It is reasonably well established that preschoolers are much less active rehearsers than are older children (see, e.g., Flavell, 1970), and such simple rehearsal may have been further depressed for experimental Ss by the heavy emphasis placed on category-label production during training. That is, perhaps the experimental Ss' simple item rehearsal was partially replaced, rather than supplemented, by their production of category labels. If the above analysis is correct, then one would expect that FR of categorizable lists would be facilitated only if preschoolers were given training that emphasized both of the encoding activities just described. Also, one might expect that the training program employed in these experiments might be much more successful in producing the above facilitative effect with somewhat older children, since such children would be expected to have become more active in terms of simple rehearsal activities.

References

Bach, M. J., & Underwood, B. J. Developmental changes in memory attributes. Journal of Educational Psychology, 1970, 61, 292-296.

Bower, G. H., Lesgold, A. M., & Tieman, D. Grouping operations in free recall. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 481-493.

Cole, M., Frankel, F., & Sharp, D. Development of free recall learning in children. Developmental Psychology, 1971, 4, 109-123.

Entwisle, D. R. Word Associations of Young Children. Baltimore: Johns Hopkins Press, 1966.

Felzen, E., & Anisfeld, M. Semantic and phonetic relations in the false recognition of words by third- and sixth-grade children. Developmental Psychology, 1970, 3, 163-168.

Flavell, J. H. Developmental studies of mediated memory. In: H. W. Reese & L. P. Lipsitt (eds.), Advances in Child Development and Behavior, Vol. 5, N.Y.: Academic Press, 1970.

Flavell, J. H., Beach, D. H., & Chinsky, J. M. Spontaneous verbal rehearsal in a memory task as a function of age. Child Development, 1966, 37, 283-299.

Gupton, T., & Frincke, G. Imagery, mediational instructions, and noun position in free recall of noun-verb pairs. Journal of Experimental Psychology, 1970, 86, 461-462.

Hall, J. W. Errors in word recognition and discrimination by children of two age levels. Journal of Educational Psychology, 1969, 60, 144-147.

Hall, J. W. Word recognition errors by children of two age levels. Journal of Educational Psychology, 1968, 59, 420-424.

Hall, J. W., & Halperin, M. S. The development of memory-encoding processes in young children. Developmental Psychology, 1972, 6, 181.

Hall, J. W., Sekuler, R., & Cushman, W. Effects of IAR occurrence on response time during subsequent recognition. Journal of Experimental Psychology, 1969, 79, 39-42.

References (cont.)

Hall, J. W., & Ware, W. B. Implicit associative responses and false recognition by young children. Journal of Experimental Child Psychology, 1968, 6, 52-60.

Laurence, M. W. Age difference in performance and subjective organization in the free-recall learning of pictorial material. Canadian Journal of Psychology, 1966, 20, 388-399.

Moely, B. M., Olson, F. S., Halwes, T. G., & Flavell, J. B. Production deficiency in young children's clustered recall. Developmental Psychology, 1969, 1, 26-34.

Nelson, K. J. The organization of free recall by young children. Journal of Experimental Child Psychology, 1969, 8, 284-295.

Paivio, A. Imagery and Verbal Processes. New York: Holt, Rinehart, and Winston, Inc., 1971.

Palermo, D. S., & Jenkins, J. J. Oral word association norms for children in grades one through four. Research Bulletin No. 60, The Pennsylvania State University, 1966.

Piaget, J., & Inhelder, B. The Psychology of the Child. New York: Basic Books, Inc., 1969.

Underwood, B. J. Attributes of memory. Psychological Review, 1969, 76, 559-573.

Underwood, B. J. False recognition produced by implicit verbal responses. Journal of Experimental Psychology, 1965, 70, 122-129.

Vaughan, M. E. Clustering, age, and incidental learning. Journal of Experimental Child Psychology, 1968, 6, 323-334.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw Hill, 1962.